Cool Stuff at Cold Start:
BBN System for TAC 2015

Bonan Min
Marjorie Freedman
Constantine Lignos

Nov 2015

The views expressed are those of the authors and do not reflect the official policy or position of the Department of Defense or the U.S. Government. Distribution Statement “A” (Approved for Public Release. Distribution Unlimited.)

This research was developed with funding from the Defense Advanced Research Projects Agency (DARPA) DEFT Program.
Outline

• BBN’s 2014 ColdStart System
• What’s New
• Entity Discovery & Linking
• Slot Filling
• Experiments and Analysis
• Conclusion
Outline

• BBN’s 2014 ColdStart System
  • What’s New
  • Entity Discovery & Linking
  • Slot Filling
  • Experiments and Analysis
  • Conclusion
BBN’s 2014 ColdStart System

- **Finding local facts**
  - SERIF NLP: syntactic parsing, proposition, etc
  - **Mentions/entities:** SERIF
  - **Relations**
    - SERIF: ACE relations/events
    - Automatically proposed & manually filtered patterns
      - Bootstrap learning
      - Distant supervision
      - Prior-year TAC slot filling annotation
    - Manually created patterns

---

He \_PER\_ graduated from **Georgia State** \_ORG\_. Babcock \_PER\_ was born in July 1900 \_TIME\_ and grew up \_GPE\_ in Ontario.

---

- **Mention detection & coreference resolution**
- **Entity Linking**
- **Relation & Event extraction**
- **KB Construction**
  - Knowledge Base Construction
    - Consolidation
      - add inverses, cardinalities, deduplication, etc
    - Inference & Reasoning
    - Knowledge Base

---

© 2015 Raytheon BBN Technologies. All Rights Reserved.
BBN’s 2014 ColdStart System

• **Connecting local facts to an entity**
  - Identify mentions with SERIF
  - Link non-names (e.g. he) to named entities with SERIF’s within-document coreference
  - Use BBN Actor matching tool to provide corpus-level IDs
    - Matching against a cleaned version of Freebase
    - For name strings that are not matched, clustering based on textual similarity

• **Relations are also deduplicated**

```
He_PER graduated from Georgia State_Org.
Babcock_PER was born in July 1900_Time and grew up ... in Ontario_GPE.
```

```
EID1: Bob Smith
EID2: John Babcock
```

Source corpus

Fact Finding

Mention detection & coreference resolution

Entity Linking

Relation & Event extraction

KB Construction

Knowledge Base Construction

Consolidation
add inverses, cardinalities, deduplication, etc

Inference & Reasoning

Knowledge Base

Inference & Reasoning
Outline

• BBN’s 2014 ColdStart System
• What’s New
  • Entity Discovery & Linking
  • Slot Filling
  • Experiments and Analysis
• Conclusion
What’s New: Outline

• Entity Discovery and Linking
  – Address overlinking of names across documents (and within a document)
  – Improve in-doc coreference: coreference ensemble

• Slot Filling
  – Improve pattern-based extractor
    • Additional patterns proposed from new training sources (CS2014, Rich ERE)
    • Restricted generic patterns
  – Add a statistical relation extraction component
Outline

• BBN’s 2014 ColdStart System
• What’s New

• **Entity Discovery & Linking**
  – Why focused on Entities?
  – Prevent overlinking
  – Improve with-doc coreference

• Slot Filling
• Experiments and Analysis
• Conclusion
Focus on Entities: Entry Points

- No credits if missed or miss-typed the entry points

- Cold start queries have multiple entry points
  - *2015 CS adds multiple entry points: In the MAX scoring, a system can answer a query if it finds any of the entry points

1 true positive, per-query recall 1.0

1 Miss, per-query recall 0
Focus on Entities: Cross-Doc Clusters

• Both within and cross-document linking of names are imperfect

**Within Document Mistakes**

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPE (Mexico)</td>
<td><strong>NYT ENG 20131105.0194:</strong> Vicious Mexican drug cartels smuggle cocaine, ... through ... <strong>New Mexico</strong> to Atlanta ... in the United States. ... smuggle drugs,... large amounts of cash through Las Cruces, a <strong>New Mexico</strong> city northwest of El Paso</td>
</tr>
<tr>
<td>Mexico, Mexican ...</td>
<td></td>
</tr>
<tr>
<td>New Mexico</td>
<td></td>
</tr>
<tr>
<td>PER (Ehud Barak)</td>
<td>44f152b60a8d7bcf65f7b0344764d0d8:<strong>Barak</strong> endorses <strong>Barack</strong>, touts US security support...Defense Minister <strong>Ehud Barak</strong> said Monday night that Barack Obama has been the most supportive president on matters of Israeli security ...</td>
</tr>
<tr>
<td>Ehud Barak, Barack,...</td>
<td></td>
</tr>
<tr>
<td>Barack, Barack Obama</td>
<td></td>
</tr>
</tbody>
</table>

**Name Strings Assigned to a Large Cluster**

<table>
<thead>
<tr>
<th>GPE: “U.S.” 20,550 mentions</th>
</tr>
</thead>
<tbody>
<tr>
<td>United States (30%), U.S. (20%), America (15%), ..., North America (139 mentions:0.6%), South America, Central America, ... Latin American, Colorado, IDAHO</td>
</tr>
</tbody>
</table>
Outline

• BBN’s 2014 ColdStart System
• What’s New

• **Entity Discovery & Linking**
  – Why focused on Entities?
  – Prevent overlinking
  – Improve with-doc coreference

• Slot Filling
• Experiments and Analysis
• Conclusion
Improving Within Document Coreference

- Use knowledge (from the Actor DB) about what entities exist in the world to prevent overlinking

New Information

- Mention/Name tagging
- mention-level Entity Linking
- In-doc Coref
- Entity-level Entity Linking

(KB)GPE: Mexico
GPE (Mexico)
Mexico, Mexican ...
New Mexico

(KB)GPE: New Mexico
GPE (Mexico)
Mexico, Mexican ...
New Mexico

GPE (U.S.)
United States, U.S.,
American ...
Latin American

GPE (U.S.)
United States, U.S.,
American ...

PER (Ehud Barak)
Ehud Barak, Barak, ...
Barack, Barack Obama

PER (Ehud Barak)
Ehud Barak, Barak, ...

GPE (New Mexico)
New Mexico

GPE (Latin American)
Latin American

PER (Barack Obama)
Barack, Barack Obama

Expected General Improvement

© 2015 Raytheon BBN Technologies. All Rights Reserved.
Improving KB Clusters

• EL created clusters on per-document basis
• KB consolidation: improve cluster formation at corpus-level
  – identify core, high frequency names
  – Split outliers into separate entities

GPE: U.S. 20,550 mentions

(30%) United States,
(20%) U.S.
(15%) America
... North
America (139:0.6%)
South America (123)
Central America (71)
... Latin American,
Colorado, IDAHO

split “outliers”: separate doc-entities without a core name

Expected General Improvement
Optimizing for Hop-1

- Precision for hop-1 queries counts responses as incorrect if the hop-0 was incorrect
  - This makes entities involved in many relations have a high potential risk
  - In 2014, 3 queries led to >50% hop1 FPs

Find organizations that are headquartered in countries where Liz Smith resides.

1 incorrect relation in the KB:
Hop 0: 1 True Positive; 1 False Positive; Prec: 0.5
Hop 1: 1 True Positive; **1 False Positives; Prec 0.5**

1 incorrect relation in the KB:
Hop 0: 1 True Positive, 1 False Positive; Prec: 0.5
Hop 1: 1 True Positive, **8 False Positives, Prec: 0.1**
• Further split very large entities

GPE: U.S. 20550 mentions

(30%) United States, (20%) U.S. (15%) America

... North America (139:0.6%) South America (123) Central America (71) ... Latin American, Colorado, IDAHO

split core-name cluster into smaller clusters with at most N doc-entities
Results: Changes on Splitting Entities

- “BBN”: our strong base system
- Run-1: + splitting in-doc entities
- Run-2: + splitting cross-doc entities (outliers, then very large entities)
  - Improved precision at hop-0 and hop-1 level, results in BBN’s highest “All Hops” score
  - Reduces performance on mention_ceaf and B-Cube
    - optimizing for ED(L) and ColdStart may not be the same

<table>
<thead>
<tr>
<th></th>
<th>CS-SF</th>
<th></th>
<th>CS-LDC-MAX</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hop0</td>
<td>Hop1</td>
<td>All hops</td>
<td>Hop0</td>
</tr>
<tr>
<td></td>
<td>P</td>
<td>R</td>
<td>F1</td>
<td>P</td>
</tr>
<tr>
<td>BBN</td>
<td>47</td>
<td>31</td>
<td>37</td>
<td>47</td>
</tr>
<tr>
<td>Run-1</td>
<td>47</td>
<td>31</td>
<td>37</td>
<td>49</td>
</tr>
<tr>
<td>Run-2</td>
<td>50</td>
<td>28</td>
<td>36</td>
<td>52 (+3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>20 (+10)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

F1:Mention ceaf

- BBN: 71
- Run-2: 69 (-2)

F1: B-Cube

- BBN: 67
- Run-2: 66 (-1)

Examples of errors removed

CSSF15_ENG_34c343d331 org:country_of_headquarters Traditional Anglican Communion org:country_of_headquarters "Australia" 79 gpe:residents_of_country
Outline

• BBN’s 2014 ColdStart System
• What’s New
• Entity Discovery & Linking
  – Why focused on Entities?
  – Prevent overlinking
  – Improve with-doc coreference
• Slot Filling
• Experiments and Analysis
• Conclusion
Within-Doc Coreference: Augmentation

- Improvements to within document coreferences using the Actor DB are limited to name-name coref
- Coreference of nominals/pronouns to names also leads to errors
- Attempt to merge coref decisions from two systems (+Stanford)
  - Align secondary system entities to SERIF entities
    - Align using canonical mention if possible
    - Otherwise to the SERIF entity it shares the most mentions with

Diagram:

- Peter Ledochowitsch
- Ledochowitsch
- he
- him
- SERIF PER(Ledochowitsch)
- Other System Entity
Results: Coreference Augmentation

<table>
<thead>
<tr>
<th></th>
<th>Hop0 P</th>
<th>Hop0 R</th>
<th>Hop0 F1</th>
<th>Hop1 P</th>
<th>Hop1 R</th>
<th>Hop1 F1</th>
<th>All hops P</th>
<th>All hops R</th>
<th>All hops F1</th>
</tr>
</thead>
<tbody>
<tr>
<td>BBN1</td>
<td>CSSF</td>
<td>46</td>
<td>31</td>
<td>37</td>
<td>12</td>
<td>13</td>
<td>13</td>
<td>29</td>
<td>25</td>
</tr>
<tr>
<td>BBN1</td>
<td>LDC-MAX</td>
<td>49</td>
<td>37</td>
<td>42</td>
<td>10</td>
<td>18</td>
<td>13</td>
<td>27</td>
<td>30</td>
</tr>
<tr>
<td>BBN3</td>
<td>CSSF</td>
<td>43</td>
<td>31</td>
<td>36</td>
<td>12</td>
<td>13</td>
<td>13</td>
<td>28</td>
<td>25</td>
</tr>
<tr>
<td>BBN3</td>
<td>LDC-MAX</td>
<td>42</td>
<td>37</td>
<td>40</td>
<td>10</td>
<td>18</td>
<td>12</td>
<td>25</td>
<td>31</td>
</tr>
</tbody>
</table>

- **24% of added mentions are pronouns**
  - Many are first person pronouns from inside quotes
- **Many useful added mentions are from correctly detecting partial name matches**
  - *Boston Common/the Common*
- **Systems often disagree about desired extents:**
  - *Netherlands/the Netherlands, Boston/Boston’s*
  - *[President][Barack Obama]/[President Barack Obama]*
- **Over-linking problems often propagate**
  - Stanford system often links “X University” and “Y University” together, transferring the error to Serif entity
Outline

• BBN’s 2014 ColdStart System
• What’s New
• Entity Discovery & Linking
• Slot Filling
  – Improve pattern-based extractor
  – Statistical relation extraction
• Experiments and Analysis
• Conclusion
Pattern-based Extractor

- **Proposition patterns**: compact, flexible, and accurate representation of many surface-level constructions

  - John accused Mary
  - John, a friend of Sheila, accused Mary
  - Mary was accused by John
  - John accused his British friend, Mary
  - John was sorry to accuse his friend Mary

- **Cold Start patterns**
  - Regex: A1 works for A2
  - Proposition

  - verb: work*
  - subj. for

  - Bonan Min
  - BBN

Automatic pattern proposing
Pattern-based Extractor: Improvements

• Improvement #1: additional patterns proposed from Rich ERE & 2014 CS assessment

• Improvement #2: use broad-coverage patterns, e.g., “A1 of A2”
  – per:resident_of: a longtime resident of San Diego
    • A1 of A2 ^ A1:PER ^ A2:GPE ^ contain_word(A1, “resident”)
  – per:origin: an immigrant of Mexican origin
    • A1 of A2 ^ A1:PER ^ A2:Country ^ text_after_A2 (“origin”)
  – per:employee_of: president of Harvard
    • A1 of A2 ^ A1:PER ^ A2:ORG ^ contain_word(A1, “president”)

• Type of restrictions (automatically collected from context of patterns)
  – argument type, words, text before/after arguments, etc

• Apply to broad-coverage patterns
Results: Improved Pattern-based Extractor

- **BBN 2014**: BBN’s 2014 Cold Start system
- **BBN1**: higher precision and recall
  - Added more patterns proposed from the new datasets
  - Added restricted generic patterns

### CS-SF score on 2015 assessment

<table>
<thead>
<tr>
<th></th>
<th>Hop0</th>
<th></th>
<th>Hop1</th>
<th></th>
<th>All hops</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P</td>
<td>R</td>
<td>F1</td>
<td>P</td>
<td>R</td>
<td>F1</td>
</tr>
<tr>
<td>BBN2014</td>
<td>45</td>
<td>29</td>
<td>35</td>
<td>10</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>28</td>
<td>22</td>
<td>25</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BBN1</td>
<td>47(+2)</td>
<td>31(+2)</td>
<td>37(+2)</td>
<td>12</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>24</td>
<td>27</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### CS-LDC-MAX score on 2015 assessment

<table>
<thead>
<tr>
<th></th>
<th>Hop0</th>
<th></th>
<th>Hop1</th>
<th></th>
<th>All hops</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P</td>
<td>R</td>
<td>F1</td>
<td>P</td>
<td>R</td>
<td>F1</td>
</tr>
<tr>
<td>BBN2014</td>
<td>47</td>
<td>34</td>
<td>40</td>
<td>9</td>
<td>15</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>28</td>
<td>27</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BBN1</td>
<td>49(+2)</td>
<td>37(+3)</td>
<td>42(+2)</td>
<td>10</td>
<td>18</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>27</td>
<td>30</td>
<td>28</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Outline

• BBN’s 2014 ColdStart System
• What’s New
• Entity Discovery & Linking
• Slot Filling
  – Improve pattern-based extractor
  – Statistical relation extraction
• Experiments and Analysis
• Conclusion
Statistical Relation Extraction

• **Training source**
  – Distant supervision (DS)
    • Freebase pairs → Gigaword
  – TAC: query → mention in reference doc
  – Rich ERE: mention-level annotation

<table>
<thead>
<tr>
<th>Dataset</th>
<th># Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>SF2013</td>
<td>5268</td>
</tr>
<tr>
<td>SF2012</td>
<td>3825</td>
</tr>
<tr>
<td>SF2014</td>
<td>3549</td>
</tr>
<tr>
<td>CS2013</td>
<td>1689</td>
</tr>
<tr>
<td>CS2014</td>
<td>1350</td>
</tr>
</tbody>
</table>

• **Features**

<table>
<thead>
<tr>
<th>Feature</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regex+ArgType</td>
<td><em>PER:smith_work</em> for <em>ORG:company</em></td>
</tr>
<tr>
<td>Prop+ArgType</td>
<td>verb:works[PER:smith:&lt;sub&gt;][ORG:company:for]</td>
</tr>
<tr>
<td>Head words</td>
<td>smith; company; smith-company</td>
</tr>
<tr>
<td>Bag of words</td>
<td>A1: John; smith A2: company between: works before/after: Amazon</td>
</tr>
<tr>
<td>Entity Type</td>
<td>A1:PER; A2:ORG; A1-A2-PER-ORG</td>
</tr>
</tbody>
</table>

Features for “John Smith works for the eCommerce company Amazon.”
Statistical Relation Extraction

- Train MaxEnt models on relation mentions

<table>
<thead>
<tr>
<th>ERE</th>
<th>TAC</th>
<th>DS</th>
</tr>
</thead>
<tbody>
<tr>
<td>per_employee_or_member_of</td>
<td>896</td>
<td>1101</td>
</tr>
<tr>
<td>per_religion_or_origin</td>
<td>368</td>
<td>595</td>
</tr>
<tr>
<td>per_place_of_residence</td>
<td>317</td>
<td>579</td>
</tr>
<tr>
<td>per_title</td>
<td>279</td>
<td>257</td>
</tr>
<tr>
<td>org_parents</td>
<td>209</td>
<td>155</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

# of positive examples in the training datasets.
*went through feature-vector deduplication
## Results: Statistical Relation Extraction

<table>
<thead>
<tr>
<th>Slot</th>
<th>BBN1</th>
<th>BBN4</th>
<th>Statistical RE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td># added</td>
<td>% added</td>
<td># added</td>
</tr>
<tr>
<td>ALL</td>
<td>236264</td>
<td>326130</td>
<td>73835</td>
</tr>
<tr>
<td>per:employee_or_member_of</td>
<td>39268</td>
<td>46488</td>
<td>7220</td>
</tr>
<tr>
<td>org:subsidiaries</td>
<td>7931</td>
<td>11902</td>
<td>3971</td>
</tr>
<tr>
<td>per:countries_of_residence</td>
<td>6428</td>
<td>17485</td>
<td>5109</td>
</tr>
<tr>
<td>per:statesorprovinces_of_residence</td>
<td>2901</td>
<td>8136</td>
<td>1386</td>
</tr>
<tr>
<td>per:cities_of_residence</td>
<td>7477</td>
<td>10503</td>
<td>3026</td>
</tr>
<tr>
<td>org:country_of_headquarters</td>
<td>2820</td>
<td>6662</td>
<td>671</td>
</tr>
<tr>
<td>org:stateorprovince_of_headquarters</td>
<td>1162</td>
<td>3936</td>
<td>292</td>
</tr>
<tr>
<td>org:city_of_headquarters</td>
<td>4428</td>
<td>4993</td>
<td>565</td>
</tr>
<tr>
<td>per:schools_attended</td>
<td>1336</td>
<td>3161</td>
<td>1825</td>
</tr>
<tr>
<td>per:stateorprovince_of_death</td>
<td>59</td>
<td>147</td>
<td>7</td>
</tr>
</tbody>
</table>

Relations added by the statistical relation models

Added Trained models (BBN4) into the base system (BBN1)
Results: Statistical Relation Extraction

- Trained model with each datasets (evaluated on CS2015 & CS2014)

<table>
<thead>
<tr>
<th>Hop0</th>
<th>P</th>
<th>R</th>
<th>F1</th>
</tr>
</thead>
<tbody>
<tr>
<td>DS</td>
<td>28</td>
<td>9</td>
<td>14</td>
</tr>
<tr>
<td>ERE</td>
<td>43</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>TAC</td>
<td>51</td>
<td>2</td>
<td>5</td>
</tr>
</tbody>
</table>

Trained model with each dataset only.
High quality features, CS 2015, STRING_CASE, CS-SF score

<table>
<thead>
<tr>
<th>Hop0</th>
<th>P</th>
<th>R</th>
<th>F1</th>
</tr>
</thead>
<tbody>
<tr>
<td>DS</td>
<td>46</td>
<td>16</td>
<td>23</td>
</tr>
<tr>
<td>ERE</td>
<td>52</td>
<td>6</td>
<td>11</td>
</tr>
<tr>
<td>TAC</td>
<td>62</td>
<td>7</td>
<td>12</td>
</tr>
</tbody>
</table>

Trained model with each dataset only.
High quality features, CS 2014

- More fine-grained features helped; but with cost of precision)

<table>
<thead>
<tr>
<th>Feature</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regex+ArgType</td>
<td>\textit{PER:smith_work_for_ORG:company}</td>
</tr>
<tr>
<td>Prop+ArgType</td>
<td>\textit{verb:works{PER:smith:sub}{ORG:company:for}}</td>
</tr>
<tr>
<td>Head words</td>
<td>\textit{smith; company; smith-company}</td>
</tr>
<tr>
<td>Bag of words</td>
<td>\textit{A1: John; A2: company \textit{between: works before/after: Amazon}}</td>
</tr>
<tr>
<td>Entity Type</td>
<td>\textit{A1:PER; A2:ORG; A1-A2-PER-ORG}</td>
</tr>
</tbody>
</table>

Features for \textit{John Smith works for the eCommerce company Amazon.}

- Trained model with datasets; also Added low-quality features, CS 2014

<table>
<thead>
<tr>
<th>Hop0</th>
<th>P</th>
<th>R</th>
<th>F1</th>
</tr>
</thead>
<tbody>
<tr>
<td>TAC</td>
<td>61 *</td>
<td>19</td>
<td>29</td>
</tr>
<tr>
<td>TAC+ERE</td>
<td>56 *</td>
<td>27</td>
<td>36</td>
</tr>
</tbody>
</table>

* Precision numbers are not quite useful since this is post-assessment evaluation
Outline

• BBN’s 2014 ColdStart System
• What’s New
• Entity Discovery & Linking
• Slot Filling
• Experiments and Analysis
• Conclusion
Experiments – official runs

- BBN1: base system
- BBN2: + split of cross-doc entities
- BBN3: + split doc entities + coreference ensemble
- BBN4: + statistical RE + inference
- BBN5: + nested names

<table>
<thead>
<tr>
<th></th>
<th>BBN1</th>
<th>BBN2</th>
<th>BBN3</th>
<th>BBN4</th>
<th>BBN5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Improved patterns</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Split cross-doc entities</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Split in-doc entities</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coreference ensemble</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statistical RE</td>
<td></td>
<td></td>
<td></td>
<td>√</td>
<td></td>
</tr>
<tr>
<td>Inference</td>
<td></td>
<td></td>
<td></td>
<td>√</td>
<td></td>
</tr>
<tr>
<td>Nested names</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>√</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>CS-SF</th>
<th></th>
<th>CS-LDC-MAX</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hop0</td>
<td>Hop1</td>
<td>All hops</td>
<td>Hop0</td>
</tr>
<tr>
<td></td>
<td>P</td>
<td>R</td>
<td>F1</td>
<td>P</td>
</tr>
<tr>
<td>BBN1</td>
<td>46</td>
<td>31</td>
<td>37</td>
<td>12</td>
</tr>
<tr>
<td>BBN2</td>
<td>49</td>
<td>28</td>
<td>36</td>
<td>23</td>
</tr>
<tr>
<td>BBN3</td>
<td>43</td>
<td>31</td>
<td>36</td>
<td>12</td>
</tr>
<tr>
<td>BBN4</td>
<td>37</td>
<td>36</td>
<td>36</td>
<td>12</td>
</tr>
<tr>
<td>BBN5</td>
<td>45</td>
<td>29</td>
<td>36</td>
<td>12</td>
</tr>
</tbody>
</table>

Official scores on the preliminary assessments
Experiments: What Helped

- BBN1: base system (BBN 2014 + more training data + restricted generic patterns)
- BBN2: BBN1 + split of cross-doc entities
- BBN3: BBN1 + split doc entities + coreference ensemble
- BBN4: BBN1 + statistical RE + inference
- BBN5: BBN1 + with finding nested names e.g., identify "US" in mention “US Army”.

Scores on the preliminary assessments. * BBN Running Scoring (not NIST reported Scores); use post-hoc flag for all
# Overall Stats

<table>
<thead>
<tr>
<th></th>
<th>#entities</th>
<th>PER</th>
<th>ORG</th>
<th>GPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>BBN1</td>
<td>210801</td>
<td>124700</td>
<td>67966</td>
<td>18135</td>
</tr>
<tr>
<td>BBN2</td>
<td>217885 (+3.3%)</td>
<td>125031</td>
<td>69525</td>
<td>23329 (+28.6%)</td>
</tr>
<tr>
<td>BBN3</td>
<td>210812</td>
<td>124702</td>
<td>67974</td>
<td>18136</td>
</tr>
<tr>
<td>BBN4</td>
<td>210801</td>
<td>124700</td>
<td>67966</td>
<td>18135</td>
</tr>
<tr>
<td>BBN5</td>
<td>214298 (+1.6%)</td>
<td>127771(+2.4%)</td>
<td>67213</td>
<td>19314(6.5%)</td>
</tr>
</tbody>
</table>

## Distribution of slots in BBN1

- **per:title** 19%
- **org:employees_or_members** 17%
- **org:top_members_employees** 14%
- **org:subsidiaries** 14%
- **gpe:residents_of_city** 13%
- **per:countries_of_residence** 13%
- **org:city_of_headquarters** 13%
- **per:statesorprovinces_of_residence** 13%
- **org:country_of_headquarters** 13%
- **per:origin** 13%
- **org:alternate_names** 13%
- **per:parents** 13%
- **org:students** 13%
- **org:stateorprovince_of_headquarters** 13%
- **per:organizationsFounded** 13%
- **per:employee_or_member_of** 13%
- **per:top_member_employee_of** 13%
- **org:parents** 13%
- **per:cities_of_residence** 13%
- **gpe:employees_or_members** 13%
- **gpe:residents_of_country** 13%
- **org:alternate_names** 13%
- **gpe:headquarters_in_city** 13%
- **gpe:residents_of_stateorprovince** 13%
- **gpe:headquarters_in_country** 13%
- **per:spouse** 13%
- **per:children** 13%
- **per:schools_attended** 13%
- **org:founded_by** 13%
- **gpe:headquarters_in_stateorprovince** 13%
- **per:date_of_death** 13%
Conclusion: Lessons Learnt

• Prevent overlinking of entities helped
• More and better pattern improves recall (and precision)
  – More patterns from additional training data
  – Restricted generic patterns
• Statistical relation extraction boost recall at the cost of precision
Thanks!